Estabelecimento de um Pacote de Ensaio de Proficiência para avaliação de laboratórios em análise de produtos de software

Sérgio Câmara, Thais Barras, Wilson Melo, Wladmir Chapetta, Raphael Machado*

[1]Instituto Nacional de Metrologia, Qualidade e Tecnologia (Inmetro),
Av. Nossa Senhora das Graças, 50, Xerém, Duque de Caxias, 25250-020, RJ – Brasil

{smcamara,tmbarras,wsjunior,wachapetta,rcmachado}@inmetro.gov.br

Abstract. *Laboratory accreditation is one of the solutions to fill the lack of resources in the area of software product analysis. Proficiency Test is a mechanism used by a legal authority to evaluate the competence of laboratories in a given subject. In this work, we propose the establishment of a Proficiency Test Package to guide the evaluated laboratories in the analysis of software products, defining necessary and auxiliary items in order to execute a round of evaluation. Finally, we describe the procedure to guarantee the delivery of the package in a fair, secure, and authenticated way, and with receipt control.*

Resumo. *A acreditação de laboratórios é uma das soluções para suprir a falta de recursos na área de análise de produto de software. O Ensaio de Proficiência é um mecanismo utilizado por uma autoridade legal para avaliar a competência de laboratórios em determinado assunto. Neste trabalho, propomos o estabelecimento de um Pacote de Ensaio de Proficiência a fim de orientar os laboratórios avaliados na análise de produtos de software, definindo seus itens necessários e auxiliares para execução de uma rodada de avaliação. Por fim, descrevemos o procedimento para garantir a entrega do pacote de EP de maneira justa, segura, autenticada, e com controle de recebimento.*

1. Introdução

Determinados produtos de software estão sujeitos à avaliação da conformidade por uma autoridade legal anterior à sua introdução no mercado. Dispositivos como medidores de energia elétrica, relógio eletrônico de ponto, balança, entre outros, são exemplos de produtos que necessitam serem avaliados de acordo uma série de requisitos (técnicos e de qualidade) obrigatórios para que tenha seu modelo aprovado. Uma das etapas da aprovação de modelo consiste justamente na inspeção do software contido nesses dispositivos, a fim de garantir o seu correto funcionamento.

Com o crescimento do número de produtos de software e de dispositivos com sistema embarcado, há cada vez mais a necessidade de recursos e esforço empregado por parte da autoridade legal a fim de realizar as devidas análises. Como uma solução viável e escalável para essa demanda crescente, laboratórios independentes podem ser contratados para realizar tal função, desempenhando os procedimentos estipulados pela autoridade legal. No Brasil, esses laboratórios, portanto, devem ser acreditados em escopos associados

*Trabalho apoiado por CNPq, Faperj e Projeto SHCDCiber.

à norma ISO/IEC 17025 [ISO 2017], além de ser desejável, também, que eles sejam capazes de demonstrar proficiência em análise de software. Para isso, o mecanismo de Ensaio de Proficiência (EP) [ISO 2010] pode ser utilizado para avaliar o desempenho de um conjunto de laboratórios participantes com base em critérios pré-estabelecidos, validando os métodos executados e avaliando suas competências técnicas no assunto. É de responsabilidade da autoridade legal (o "provedor") realizar as rodadas de EP.

Uma vez estabelecida a metodologia que será utilizada para avaliar os laboratórios em uma rodada de EP, é necessário, também, estabelecer o que deverá ser disponibilizado aos laboratórios para que tal rodada seja viabilizada. Neste trabalho, propomos o estabelecimento de um Pacote de Ensaio de Proficiência a fim de orientar as atividades dos laboratórios avaliados em relação ao Ensaio de Proficiência em análise de produtos de software. Esse pacote reúne todos os arquivos necessários e auxiliares (como a documentação, especificação de testes, ambiente de testes, avaliação e métricas, ferramentas sugeridas, etc) para execução dos testes do Ensaio de proficiência. Descrevemos, também, como o pacote é disponibilizado aos laboratórios participantes, de forma a estabelecer uma entrega segura, autenticada, justa e com controle de recebimento por parte do provedor da rodada.

A metodologia aqui considerada para avaliar os laboratórios é baseada na cobertura de código-fonte, na qual um laboratório deverá conseguir executar as mesmas linhas de código de um determinado software a fim de reproduzir testes de unidade semelhantes aos que foram estipulados pelo provedor da rodada (mais detalhes na Seção 3).

2. Trabalhos Relacionados

Dentre os programas internacionais para acreditação de laboratórios em avaliação de produtos de software, destacamos, nesta seção, os esforços dos Estados Unidos e da França.

O *National Voluntary Laboratory Accreditation Program* (NVLAP) é o programa americano de acreditação de laboratórios administrado pelo *National Institute of Standards and Technology* americano (NIST) [NIST 2017]. Através dos *Laboratory Accreditation Programs* (LAPs), o NVLAP estabelece, desenvolve e implementa ações para credenciar laboratórios de ensaio e calibração que são considerados competentes para realizar tais atividades. Dentre os LAPs desenvolvidos, dois se relacionam com tema deste trabalho: o *Common Criteria Testing* LAP, no qual o EP para o escopo dessa acreditação utiliza a avaliação inicial, porém não existe um item de ensaio e nenhum procedimento específico de ensaio; e o *Cryptographic & Security Testing* LAP, o qual não relata programa para EP de laboratórios credenciados até o momento deste trabalho.

O *Certification de Sécurité de Premier Niveau* (CSNP) [ANSSI 2015] é uma certificação francesa de segurança de primeiro nível para produtos de tecnologia da informação oferecida pela *Autorité Nationale en matière de Sécurité et de défense des Systèmes d'Information* (ANSSI). A ANSSI licencia laboratórios externos, validando suas habilidades em análise técnica de segurança para avaliação de conformidade de produtos de TI. Até o momento, a ANSSI não possui um ensaio de proficiência formal para avaliar os laboratórios licenciados.

3. Ensaio de proficiência por cobertura de código

A cobertura de código é uma métrica utilizada em análise de teste de software que descreve a quantidade de código-fonte que foi testada [Diego Torres Milano 2011]. Para essa metodologia de Ensaio de Proficiência, o conceito de cobertura de código e testes de unidade são utilizados de maneira diferente da abordada em Engenharia de software. Nessa metodologia, é definido que, dado um código fonte de um software (item de EP) e dado um conjunto de linhas cobertas desse código (relatório de referência) por um teste de unidade de referência implementado pelo provedor da rodada, o laboratório participante deverá tentar cobrir o mesmo conjunto de linhas, sem conhecer o teste de unidade de referência. A metodologia de EP por cobertura de código parte do princípio que a capacidade de analisar códigos fonte de um software é uma competência requerida para laboratórios envolvidos com análise de software.

O provedor poderá sugerir mais de um teste por rodada de EP, apresentando aos participantes apenas os relatórios de referências necessários para cada teste, e não seus testes de unidade implementados. O laboratório deverá, então, implementar seu próprio teste de unidade e emitir seu relatório, contendo o conjunto de linhas de código por ele coberto. O laboratório participante deverá entregar tanto seus relatórios quanto seus testes de unidades para serem avaliados pelo provedor da rodada. O provedor, por sua vez, calcula uma nota em relação à proficiência de cada laboratório de acordo com métricas de avaliação pré-estabelecidas. Caso o laboratório obtenha uma boa avaliação, pode-se concluir ele tem um bom entendimento do código relacionado àqueles testes e, por consequência, que o laboratório é proficiente na atividade de análise de software.

4. Pacote de Ensaio de Proficiência proposto

O Pacote de Ensaio de Proficiência é um pacote digital que inclui os arquivos necessários e auxiliares para a realização de uma rodada do Ensaio de Proficiência. O Pacote de EP aqui proposto consiste de um único arquivo: uma imagem de uma máquina virtual (VM, *virtual machine*). A partir dessa imagem, o laboratório é capaz de executar um ambiente de testes virtualizado já preparado e configurado para uso no ensaio de proficiência. Incluídos no sistema de arquivos do ambiente de testes, estão os seguintes itens:

- Documentação: descrição das ferramentas recomendadas, manual para instruções de instalações e uso, descrição do processo de avaliação, requisitos, etc.
- Testes: Especificação dos testes a serem reproduzidos, biblioteca de cobertura de código, e relatórios de referência.
- Item de ensaio de proficiência: o código-fonte e uma versão funcional do produto de software escolhido para a rodada de EP.
- Ferramentas auxiliares: ambiente de desenvolvimento (IDE), compiladores e interpretadores necessários, e *scripts* para automação de atividades.

O laboratório participante poderá optar utilizar, ou não, o ambiente de testes virtualizado disponibilizado pelo provedor da rodada. Caso o participante queira utilizar seus próprios meios e ferramentas para completar os testes requeridos, ele poderá copiar, do ambiente de testes disponível, as informações sobre os Testes e sobre o Item de EP para o ambiente de sua preferência.

4.1. Definição dos itens do Pacote de EP

Ambiente de testes

O ambiente de testes sugerido consiste de uma máquina virtual com o sistema operacional Lubuntu 18.04 64-bit instalado. O Lubuntu é aqui indicado por se tratar de um sistema "leve" e que não demanda grandes recursos computacionais da máquina host para sua execução. Nos nossos testes, uma máquina host com processador Intel Core i5 e memória RAM de 6 GB consegue hospedar esse ambiente sem problemas de desempenho.

A imagem da máquina virtual é disponibilizada através de um arquivo ".ova" (*Open Virtualization Appliance*), o qual é um padrão aberto de empacotamento e distribuição de máquinas virtuais e imagens de discos virtuais (p.ex., VMDK) compatível com diversos softwares de virtualização.

Item de Ensaio de Proficiência

O item de EP sugerido para compor o pacote é o software *Alliance P2P* [Alliance P2P]. O *Alliance* é uma rede *peer-to-peer* designada para compartilhar arquivos e estabelecer comunicação entre pessoas que se conhecem ("amigos"). O software é gratuito, foi desenvolvido em linguagem Java 5.0 e é descrito pelo desenvolvedor como um ambiente privado e seguro. O Alliance é composto por dois subsistemas independentes: o *User Interface* (UI), parte responsável por manipular a interface gráfica; e o *Core*, parte responsável pelo gerenciamento do restante do sistema. O subsistema *Core* é dividido em três pacotes: o *Comm*, que contém as classes responsáveis pelo fluxo de dados de rede; o *Node*, que é formado por classes com informações dos amigos; e o *File*, que possui classes onde é feito o gerenciamento de arquivos compartilhados.

Para a rodada de EP, as classes são selecionadas obedecendo a premissa que a classe escolhida deve pertencer a um dos pacotes do subsistema *Core* (*Comm*, *File* e *Node*). Além disso, todos os pacotes devem ter ao menos uma classe testada.

Testes

Existem muitas ferramentas de análise de cobertura de código, porém neste trabalho indicamos a biblioteca JaCoCo (*Java Code Coverage*) [Marc Hoffmann *et al.*]. A partir da instrumentação do código do software sendo testado, a biblioteca JaCoCo é capaz de medir a cobertura por linha e ramificações do código abrangido pelo caso de teste de unidade em execução. Além disso, ela emite relatórios de resultado com a percentagem de código que foi coberto em três tipos de extensões de arquivo: HTML, XML e CSV. O relatório em XML mostra em detalhes as linhas cobertas e não-cobertas para o teste em execução.

Os relatórios de referência gerados pelo provedor da rodada do EP estão no formato XML do JaCoCo. Ou seja, para cada teste do EP, existe um relatório JaCoCo como referência, detalhando quais linhas foram cobertas e quais não foram, entre todas as classes pertencentes ao software *Alliance*. Os relatórios de referência estão armazenados em um diretório da máquina virtual (ambiente de testes) disponibilizada, juntamente com a especificação de cada teste do ensaio a ser reproduzido pelo laboratório participante.

Ferramentas auxiliares

Algumas ferramentas auxiliares também são disponibilizadas no ambiente de testes para ajudar na execução dos testes do EP. Por se tratar de um IDE (ambiente de desenvolvimento integrado) gratuito e completo, o software Eclipse é indicado para utilização no EP

e se encontra instalado no ambiente de testes disponibilizado. Apesar de todos os testes do EP poderem ser implementados de maneira manual, sem depender de um ambiente de desenvolvimento específico, o Eclipse apresenta algumas facilidades para execução do ensaio. Em destaque, ele permite a criação automatizada de testes de unidade com o JUnit, além de já ter integrada nativamente a biblioteca JaCoCo, facilitando a escrita dos testes de unidade e emissão de relatórios de cobertura.

Sugerimos, também, a inclusão no pacote de EP de um script[1] de comparação entre relatórios JaCoCo. A partir desse script é possível destacar de forma automatizada as linhas que precisam ser cobertas e as linhas que não precisam ser cobertas pelos testes de unidade criados pelo laboratório. Além disso, o script é capaz de calcular as métricas de avaliação em relação a cada teste do EP. O script poderá ser reexecutado diversas vezes durante toda a implementação dos testes de unidade para que se acompanhe quais linhas restam cobrir e qual a sua nota obtida pelas métricas de avaliação em determinado momento. Caso o script não seja utilizado, o laboratório avaliado deverá buscar manualmente pelo extenso relatório JaCoCo a procura das linhas a serem cobertas/não-cobertas.

4.2. Medidas de segurança para o Pacote de EP

Nesta subseção, abordaremos as medidas de segurança necessárias à distribuição do pacote de EP para os laboratórios participantes. A seguir, descrevemos o procedimento para garantir a entrega do pacote de EP de maneira segura, autenticada, justa para todos os participantes, e com controle de recebimento por parte do provedor da rodada.

1. Na ficha de inscrição (formulário no site do provedor do EP), o laboratório participante informa uma chave pública, PU_{lab}, para validação dos resultados por ele gerados (ela tem o valor de um certificado digital auto assinado).
2. Ao final da montagem do pacote de EP, P_{EP}, o provedor da rodada deverá realizar as seguintes operações: Dada uma função de hash criptográfica, $H()$, considerada segura atualmente (p.ex., SHA-256), o pacote deverá ser assinado digitalmente pelo provedor, $SIG_{PEP} = E(PR_{prov}, h_{PEP})$, onde $E()$ é um algoritmo criptográfico seguro atualmente (p.ex., RSA-2048); PR_{prov} é a chave privada do provedor do ensaio; e $h_{PEP} = H(P_{EP})$ é valor do hash do pacote. Além disso, o provedor deve solicitar um carimbo de tempo autenticado por uma Autoridade de Carimbo de Tempo (ACT), $T_{PEP} = E(PR_{ACT}, h_{PEP}||timestamp)$, onde PR_{ACT} é a chave privada da ACT; e $timestamp$ são as informações de hora e data concatenadas com o valor de h_{PEP}.
3. Anteriormente (p.ex., 15 dias) a data estipulada para início da rodada de EP, o provedor disponibiliza o pacote de EP encriptado, P_{enc}, com uma chave secreta simétrica, K (revelada posteriormente), tal que $P_{enc} = E(K, P_{EP})$. O provedor também deve informar o valor de hash, h_{Penc}, do pacote encriptado; a assinatura digital, SIG_{PEP}; e o carimbo de tempo, T_{PEP}, correspondente ao pacote em claro.
4. Os laboratórios podem, se assim quiserem, fazer *download* do pacote com antecedência e verificar sua integridade a partir do valor de hash h_{Penc}. Esse passo mitiga problemas relacionados a eventuais indisponibilidades do serviço de *download*, ou de algum participante ser favorecido por ter uma taxa de *download* mais rápida.
5. No dia e hora exatos de início da rodada de EP, o provedor disponibiliza em seu site a chave secreta K necessária para decriptar o pacote de EP.

[1] `https://github.com/smcamara/ep/blob/master/comparaRelatorios.py`

6. Cada laboratório participante deve, após decriptar o pacote, verificar sua assinatura digital e sua estampa de tempo. O participante deve, também, informar ao provedor (p.ex., por e-mail) o valor de hash do pacote de EP, h_{PEP}, assinado digitalmente pelo próprio laboratório (verificável pela sua chave-pública). Tal evidência prova que o laboratório teve acesso ao conteúdo correto do pacote de EP.
7. Se o laboratório informar um valor de hash correto, igual a h_{PEP}, o provedor confirma o recebimento por e-mail. Caso o laboratório informar um valor de hash incorreto, o provedor notifica o mesmo que seu pacote de EP está inconsistente. O laboratório deve, então, repetir os procedimentos 4 e 6 até apresentar o valor de hash esperado, ou do contrário sua participação na rodada de EP será invalidada.

5. Conclusão

Devido ao crescimento do número de produtos de software e de dispositivos com sistema embarcado, a acreditação de laboratórios independentes tornou-se uma solução viável e escalável para suprir a necessidade de mais recursos na área de análise de software. Para isso, a competência do laboratório no assunto também deve ser demonstrada através de Ensaios de Proficiência organizados pela autoridade legal. Além da metodologia utilizada nos testes do EP, é necessário, também, estabelecer o que deve ser disponibilizado aos laboratórios para que tal rodada seja viabilizada.

O presente trabalho propôs o estabelecimento de um Pacote de Ensaio de Proficiência a fim de orientar as atividades dos laboratórios avaliados no Ensaio de Proficiência em análise de produtos de software. Definimos os itens necessários e auxiliares desse pacote, assim como detalhamos algumas escolhas sobre ferramentas e bibliotecas específicas sugeridas para a metodologia de avaliação através de cobertura de código. Por fim, descrevemos o procedimento de como o pacote de EP é disponibilizado aos laboratórios participantes de forma a garantir uma entrega segura, autenticada, justa para todos os participantes, e com controle de recebimento do pacote por parte do provedor.

Referências

Alliance P2P. `http://alliancep2p.sourceforge.net/`. [acesso: 01-07-19].

ANSSI (2015). *Licensing of evaluation facilities for the first level security certification*. 1.2 edition.

Diego Torres Milano (2011). *Android Application Testing Guide*. Packt Publishing.

ISO (2010). *ISO/IEC 17043:2010 Conformity assessment – General requirements for proficiency testing*. International Standard Organization, 1st edition.

ISO (2017). *ISO/IEC 17025:2017 General requirements for the competence of testing and calibration laboratories*. International Standard Organization, 3rd edition.

Marc Hoffmann *et al.* JaCoCo Java Code Coverage Library. `https://www.eclemma.org/jacoco/`. [acesso: 27-06-2019].

NIST (2017). National Voluntary Laboratory Accreditation Program (NVLAP). `https://www.nist.gov/nvlap`. [acesso: 27-06-2019].